[53] Seidenfeld, T. (1985). Calibration, coherence, and scoring rules. *Phil. Sci.*, **52**, 274–294.

[54] Singpurwalla, N. D. and Song, M. S. (1987). The analysis of Weibull lifetime data incorporating expert opinion. In *Probability and Bayesian Statistics*, R. Viertl, ed. Plenum Press, New York, pp. 431–442.

[55] Spetzler, C. S. and Staël von Holstein, C. S. (1975). Probability encoding in decision analysis. *Management Sci.*, **22**, 340–358.

[56] Spiegelhalter, D. J. and Freedman, L. S. (1986). A predictive approach to selecting the size of a clinical trial, based on subjective clinical opinion. *Statist. Med.*, **5**, 1–13.

[57] Tversky, A. (1974). Assessing uncertainty. *J. R. Statist. Soc. B*, **36**, 148–159.

[58] Tversky, A. and Kahneman, D. (1974). Judgment under uncertainty: heuristics and biases. *Science*, **185**, 1124–1131.

[59] Vardeman, S. and Meeden, G. (1983). Calibration, sufficiency and domination considerations for Bayesian probability assessors. *J. Amer. Statist. Ass.*, **78**, 808–816.

[60] Wallsten, T. S. and Budescu, D. V. (1983). Encoding subjective probabilities: a psychological and psychometric review. *Management Sci.*, **29**, 151–173.

[61] Winkler, R. L. (1967). The assessment of prior distributions in Bayesian analysis. *J. Amer. Statist. Ass.*, **62**, 776–800.

[62] Winkler, R. L. (1967). The quantification of judgment: some methodological suggestions. *J. Amer. Statist. Ass.*, **62**, 1105–1120.

[63] Winkler, R. L. (1981). Combining probability distributions from dependent information sources. *Management Sci.*, **27**, 479–488.

[64] Winkler, R. L. (1986). On good probability appraisers. In *Bayesian Inference and Decision Techniques*, P. Goel and A. Zellner, eds. Elsevier Science, New York, pp. 265–278.

[65] Winkler, R. L. and Murphy, A. H. (1968). Good probability appraisers. *J. Appl. Meteorol.*, **7**, 751–758.

[66] von Winterfeldt, D. and Edwards, W. (1986). *Decision Analysis and Behavioral Research*. Cambridge University Press.

[67] Wolfson, L. J. (1995). Elicitation of Priors and Utilities for Bayesian Analysis. Ph.D. Thesis, Department of Statistics, Carnegie Mellon University, Pittsburgh.

[68] Wolpert, R. L. (1989). Eliciting and combining subjective judgments about uncertainty. *Int. J. Technol. Assessment Health Care*, **4**, 537–557.

[69] Wright, G. and Ayton, P., eds. (1994). *Subjective Probability*. Wiley, New York.

(BAYESIAN INFERENCE
POSTERIOR DISTRIBUTIONS
PREDICTIVE ANALYSIS
PRIOR DISTRIBUTIONS)

LARA J. WOLFSON

EMPIRICAL BIVARIATE QUANTILE-PARTITIONED DISTRIBUTION

Epidemiologists sometimes collect bivariate continuous data on a sample of individuals, calculate the empirical quantiles* of the marginal data, and then partition the original data into two-way contingency tables* with row and column categories defined by these values [16]. For example, Pietinen et al. [13, 14] conducted an extensive study on Finnish men aged 55–69 to test the reproducibility and validity of dietary measurement techniques. As part of this study, the vitamin E intake of 157 men was measured by a prospective food record diary (X) and a retrospective food use questionnaire (Y). In turn, these bivariate measurements were partitioned into categories defined by the empirical quintiles of the X and Y variables to create Table 1 [4]. The term *empirical bivariate quantile-partitioned* (EBQP) *distribution* describes the distribution of counts in such tables [3].

Because the original data are partitioned by the empirical quintiles, the marginal totals of

Table 1 EBQP Table for the Vitamin E Data Partitioned by Empirical Quintiles[a]

Food Record Quintiles	Food Use Questionnaire Quintiles 1 (low)	2	3	4	5 (high)	Total
1 (low)	13	13	4	1	0	31
2	9	9	7	4	2	31
3	7	5	7	7	6	32
4	2	2	12	9	6	31
5 (high)	0	2	2	10	18	32
Total	31	31	32	31	32	157

[a]Created from original data in a study by Pietinen et al. [13], adapted from ref. [4].

the rows and column are fixed at 31 or 32 observations, even though the interior counts of the table are still random. For example, 12 individuals fell in both the fourth quintile of the food record measurements and the third quintile of the food use questionnaire measurements.

The special case of 2×2 EBQP tables partitioned by empirical medians has received considerable study. In 1899, Sheppard [15] proposed studying the agreement between two bivariate continuous measurements by constructing such tables. Blomqvist [1] discussed a measure of agreement*, q, adapted from Mosteller [12] for such tables. This measure, called the *medial correlation* [10] or the *quadrant measure* [11], is algebraically equivalent to Kendall's tau* in 2×2 tables*. In order to study the distribution of such statistics, Blomqvist developed an exact theory and derived an asymptotic variance for 2×2 tables under certain regularity conditions, amended by Konijn [10]. Elandt [6, 7] discussed the power of Blomqvist's exact test of independence under bivariate normal alternatives.

Borkowf et al. [3] derived the asymptotic normal distribution theory for the cell proportions in $r \times c$ EBQP tables, and gave a special formula for the 2×2 case. They found, to their surprise, that Blomqvist's asymptotic result for 2×2 EBQP tables was correct only in special cases. Measures of agreement calculated from $r \times c$ EBQP tables, such as kappa* or weighted kappa, are functions of the cell proportions in these tables. Hence, the asymptotic normal distribution of these measures can be derived from the asymptotic joint normal distribution of the cell proportions, and the variances of these measures can be calculated by the delta method*.

The EBQP method of constructing tables and the corresponding method of inference for measures calculated from these tables are both nonparametric*. By contrast, if one is willing to assume that the original bivariate measurements come from a particular parametric family, one can estimate the parameters that determine the shape of the underlying distribution, calculate the parametric estimates of the population quantiles, and then use these values to estimate the expected counts in tables analogous to Table 1. Borkowf and Gail [5] developed the parametric method in order to study the asymptotic relative efficiency* of EBQP methods and to improve the precision of estimates of measures of agreement.

Here we review the asymptotic distribution theory for $r \times c$ EBQP tables, discuss the efficiency of EBQP methods compared to parametric methods, and illustrate the use of these methods with an example.

NOTATION AND ASSUMPTIONS

Let the bivariate continuous sample $\{(X_k, Y_k)\}$ $(k = 1, 2, \ldots, t)$ be i.i.d. from the distribution F. Let $F(x, y)$ have marginal distributions $G(x)$ and $H(y)$ and conditional distributions $G(x \mid y)$ and $H(y \mid x)$. Also, let $\hat{F}(x, y)$, $\hat{G}(x)$, and $\hat{H}(y)$ denote the corresponding right-continuous empirical distribution functions (*see* EDF STATISTICS).

Next, let $\{\gamma_i\}(i = 0, 1, \ldots, r)$ and $\{\eta_j\}(i = 0, 1, \ldots, c)$ denote two increasing sets of cumulative marginal proportions such that $\gamma_0 = \eta_0 = 0$ and $\gamma_r = \eta_c = 1$. For example, for quintiles, $r = c = 5$, $\gamma_i = i/5$, and $\eta_j = j/5$. In turn, define the population quantiles $\xi_i = G^{-1}(\gamma_i)$ and $\psi_j = H^{-1}(\eta_j)$, the empirical quantiles $u_i = \inf\{u : \gamma_i \leq \hat{G}(u)\}$ and $v_j = \inf\{v : \eta_j \leq \hat{H}(v)\}$, the conditional proportions $\gamma_{i|j} = G(\xi_i \mid \psi_j)$ and $\eta_{j|i} = H(\psi_j \mid \xi_i)$, and the cumulative proportions $\phi_{ij} = F(\xi_i, \psi_j)$.

It is assumed that $g(x) = G'(x)$ and $h(y) = H'(y)$ exist and are positive at the selected population quantiles, so $\xi_i = G^{-1}(\gamma_i)$ and $\psi_j = H^{-1}(\eta_j)$ are uniquely defined. It is also assumed that $F(x, y)$ is differentiable as a function of (x, y) at each (ξ_i, ψ_j).

Then, the proportion of counts in the (i, j)th category defined by $u_{i-1} < x \leq u_i$ and $v_{j-1} < y \leq v_j$ is

$$p_{ij} = \hat{F}(u_i, v_j) - \hat{F}(u_{i-1}, v_j) - \hat{F}(u_i, v_{j-1}) + \hat{F}(u_{i-1}, v_{j-1}). \quad (1)$$

Thus, the cell counts in the $r \times c$ EBQP table are given by $\{p_{ij}t\}$. As $t \to \infty$, each empirical proportion p_{ij} tends to the asymptotic

parametric methods, the corresponding point estimates and SEs are $\hat{\kappa} = 0.242 \pm 0.027$, $\hat{\kappa}_w = 0.627 \pm 0.042$, and $\hat{\alpha}_{111} = 0.551 \pm 0.031$.

Under EBQP methods, $\hat{\kappa}_w$ depends on all the cells of the table, whereas $\hat{\kappa}$ depends only on the diagonal cells and $\hat{\alpha}_{111}$ depends only on the (1,1) cell. By contrast, parametric methods use information from all of the underlying data through the parameter estimates (in this case, the sample correlation) to estimate these measures. Thus, it is not surprising that EBQP methods are most efficient for estimating κ_w and least efficient for estimating κ and α_{111}, compared to parametric methods.

DISCUSSION

The EBQP theory presented here yields results that differ from those appropriate for statistics calculated from tables with the multinomial distribution*. Such tables are obtained if the original data are partitioned by the population quantiles rather than the empirical quantiles. Fleiss et al. [9] derived the asymptotic variances of kappa and weighted kappa calculated from multinomial tables, but these variances, though correct for multinomial tables, differ from the correct variances for statistics calculated from EBQP tables [3].

As the above example illustrates, EBQP methods can be inefficient compared to parametric methods for estimating certain measures of agreement. For BVN data, EBQP estimates of κ and α_{111} are less than 41% efficient for a range of correlations and table dimensions studied in ref. [5]. By contrast, EBQP estimates of κ_w become increasingly efficient as the table dimensions increase, for moderate correlations. The increased efficiency of parametric methods, however, comes at the risk of bias from misspecifying the parametric model [5].

EBQP methods are most helpful when used in conjunction with methods that directly examine the original bivariate continuous data. A full discussion of EBQP and parametric methods in the epidemiological context appears in ref. [4]. Sample computer code for performing EBQP and parametric calculations appears in ref. [2] and, at the time of this writing, can be obtained from the first author.

Acknowledgment

This entry was prepared while C.B.B. held a National Research Council–National Institutes of Health Research Associateship at the National Heart, Lung, and Blood Institute's Office of Biostatistics Research.

References

[1] Blomqvist, N. (1950). On a measure of dependence between two random variables. *Ann. Math. Statist.*, **21**, 593–600. (First attempt to derive the asymptotic distribution of 2 × 2 contingency tables partitioned by empirical medians.)

[2] Borkowf, C. B. (1997). The Empirical and Parametric Bivariate Quantile-Partitioned Distributions. Doctoral Dissertation, Field of Statistics, Cornell University, Ithaca, New York. (Sample computer code for EBQP and parametric methods.)

[3] Borkowf, C. B., Gail, M. H., Carroll, R. J., and Gill, R. D. (1997). Analyzing bivariate continuous data grouped into categories defined by empirical quantiles of marginal distributions. *Biometrics*, **53**, 1054–1069. (Derives the asymptotic theory and inferential procedures for the EBQP distribution.)

[4] Borkowf, C. B. and Gail, M. H. (1997). On measures of agreement calculated from contingency tables with categories defined by the empirical quantiles of the marginal distributions. *Amer. J. Epidemiol.*, **146**, 520–526. (Presents EBQP and parametric methods from an epidemiological perspective.)

[5] Borkowf, C. B. and Gail, M. H. (1997). Parametric methods for bivariate quantile-partitioned data and the efficiency of corresponding nonparametric methods. *Biometrical J.*, **39**, 919–934. (Derives a parametric counterpart to the EBQP distribution and studies the asymptotic relative efficiency of EBQP methods.)

[6] Elandt, R. (1957). A non-parametric test of tendency. *Bull. Acad. Polon. Sci.*, **5**, 187–190. (Provides critical values for Blomqvist's exact test of independence.)

[7] Elandt, R. C. (1962). Exact and approximate power function of the nonparametric test of tendency. *Ann. Math. Statist.*, **33**, 471–481. (Discusses the power of Blomqvist's exact test of independence under bivariate normal alternatives.)

[8] Fleiss, J. L. and Cohen, J. (1973). The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ. and Psych. Meas.*, **33**, 613–619.

[9] Fleiss, J. L., Cohen, J., and Everitt, B. S. (1969). Large sample standard errors of kappa and weighted kappa. *Psych. Bull.*, **72**, 323–327.

proportion

$$\pi_{ij} = F(\xi_i, \psi_j) - F(\xi_{i-1}, \psi_j) - F(\xi_i, \psi_{j-1}) + F(\xi_{i-1}, \psi_{j-1}). \quad (2)$$

The empirical marginal proportions p_{i+} and p_{+j} are constant given t, while the asymptotic marginal proportions $\pi_{i+} = \gamma_i - \gamma_{i-1}$ and $\pi_{+j} = \eta_j - \eta_{j-1}$ are constant for all t.

ASYMPTOTIC THEORY

Utilizing (1), the asymptotic joint normal distribution, expectations, and covariances of the $\{p_{ij}\}$ can be derived from those of the $\{\hat{F}(u_i, v_j)\}$.

One can approximate $\hat{F}(u_i, v_j)$ in terms of $\hat{F}(\xi_i, \psi_j)$, $\hat{G}(\xi_i)$, and $\hat{H}(\psi_j)$. Define the vectors $\lambda'_{ij} = (1, -\eta_{j|i}, -\gamma_{i|j})$, $\mu'_{ij} = (0, \gamma_i, \eta_j)$, and $w'_{ij} = (\hat{F}(\xi_i, \psi_j), \hat{G}(\xi_i), \hat{H}(\psi_j))$. Then [3]

$$\hat{F}(u_i, v_j) = \lambda'_{ij}(w_{ij} - \mu_{ij}) + o_p(t^{-1/2}). \quad (3)$$

Since $t^{1/2}(\hat{F} - F, \hat{G} - G, \hat{H} - H)$ evaluated at the population quantiles has an asymptotic joint normal distribution, (3) implies that each $t^{1/2}\hat{F}(u_i, v_j)$ tends to normality, and moreover the vector $t^{1/2}\{\hat{F}(u_i, v_j)\}_{ij}$ jointly tends to normality.

Let $m = \min\{i, k\}$ and $n = \min\{j, l\}$. Then [3] for every sample size t,

$$E[\hat{F}(\xi_i, \psi_j)] = \phi_{ij}, \quad (4)$$

$$\mathrm{Cov}[t^{1/2}\hat{F}(\xi_i, \psi_j), t^{1/2}\hat{F}(\xi_k, \psi_l)] = \phi_{mn} - \phi_{ij}\phi_{kl}; \quad (5)$$

also $G(x) = F(x, \infty)$ and $H(y) = F(\infty, y)$, so $\gamma_i = \phi_{ic}$ and $\eta_j = \phi_{rj}$. Thus, (5) implies that

$$\mathrm{Cov}[t^{1/2}w_{ij}, t^{1/2}w_{kl}] = \begin{bmatrix} \phi_{mn} - \phi_{ij}\phi_{kl} & \phi_{mj} - \phi_{ij}\gamma_k & \phi_{in} - \phi_{ij}\eta_l \\ \phi_{ml} - \gamma_i\phi_{kl} & \gamma_m - \gamma_i\gamma_k & \phi_{il} - \gamma_i\eta_l \\ \phi_{kn} - \eta_j\phi_{kl} & \phi_{kj} - \eta_j\gamma_k & \eta_n - \eta_j\eta_l \end{bmatrix} \equiv \Omega_{ijkl}. \quad (6)$$

It follows from (4), (5), and (6) that $t^{1/2}[\hat{F}(u_i, v_j) - \phi_{ij}]$ and $t^{1/2}[\hat{F}(u_k, v_l) - \phi_{kl}]$ are jointly asymptotic normal with mean zero and covariance

$$\lambda'_{ij}\Omega_{ijkl}\lambda_{kl}. \quad (7)$$

In particular, the asymptotic variance of $t^{1/2}[\hat{F}(u_i, v_j) - \phi_{ij}]$ can be written without matrix notation as

$$\phi_{ij}(1 - \phi_{ij}) + \eta^2_{j|i}\gamma_i(1 - \gamma_i) + \gamma^2_{i|j}\eta_j(1 - \eta_j) - 2\eta_{j|i}\phi_{ij}(1 - \gamma_i) - 2\gamma_{i|j}\phi_{ij}(1 - \eta_j) + 2\eta_{j|i}\gamma_{i|j}(\phi_{ij} - \gamma_i\eta_j). \quad (8)$$

In 2×2 tables partitioned by empirical medians, (8) reduces to Blomqvist's result only when $\gamma_{1|1} = \eta_{1|1} = \frac{1}{2}$. This condition holds not only under independence, but also for the bivariate normal (BVN) distribution with nonzero correlation, for instance.

For most applications, the variances and covariances involve so many terms that it is essential to use matrix notation and computer calculations. For a discussion of parameter estimation and confidence interval construction, see ref. [3].

ANALYSIS OF EXAMPLE

An analysis of the vitamin E data illustrates the use of EBQP methods in epidemiology [4]. To construct Table 1, several ties in the vitamin E data were broken by adding tiny random errors to the original data, but the manner in which these ties were broken had only a small effect on the resulting table. Consider the following three measures of agreement: kappa (κ), weighted kappa (κ_w) with quadratic weights [8], and row proportions [16] (the proportion of observations that fall in specified columns of a table, given that they fall in a specified row, e.g., $\alpha_{1|1} = \pi_{11}/\pi_{1+}$). These measures and their variances can be estimated using the methods in ref. [3], which require not only the cell counts but also the original bivariate data. Using EBQP methods, the point estimates and standard errors (SEs) for these measures of agreement are $\hat{\kappa} = 0.196 \pm 0.056$, $\hat{\kappa}_w = 0.631 \pm 0.052$, and $\hat{\alpha}_{1|1} = 0.419 \pm 0.076$.

Normal probability plots and scatter plots of the log-transformed vitamin E data suggest that the original data are consistent with an underlying bivariate lognormal distribution. Hence, the log-transformed data are consistent with a BVN distribution with correlation $\hat{\rho} = 0.681$. Using

[10] Konijn, H. S. (1956). On the power of certain tests for independence in bivariate populations. *Ann. Math. Statist.*, **27**, 300–323.

[11] Kruskal, W. H. (1958). Ordinal measures of association. *J. Amer. Statist. Ass.*, **53**, 814–861.

[12] Mosteller, F. (1946). On some useful "inefficient" statistics. *Ann. Math. Statist.*, **17**, 377–408.

[13] Pietinen, P., Hartman, A. M., Haapa, E., Rasanen, L., Haapakoski, J., Palmgren, J., Albanes, D., Virtamo, J., and Huttunen, J. K. (1988). Reproducibility and validity of dietary assessment instruments. I. A self-administered food use questionnaire with a portion size picture booklet. *Amer. J. Epidemiol.*, **128**, 655–666.

[14] Pietinen, P., Hartman, A. M., Haapa, E., Rasanen, L., Haapakoski, J., Palmgren, J., Albanes, D., Virtamo, J., and Huttunen, J. K. (1988). Reproducibility and validity of dietary assessment instruments. II. A qualitative food frequency questionnaire. *Amer. J. Epidemiol.*, **128**, 667–676.

[15] Sheppard, W. F. (1899). On the application of the theory of error to cases of normal distribution and normal correlation. *Phil. Trans. R. Soc. London A*, **192**, 101–167. (Earliest reference known to the authors of a 2 × 2 contingency table partitioned by empirical medians.)

[16] Willett, W. C., Sampson, L., Stampfer, M. J., Rosner, B., Bain, C., Witschi, J., Hennekens, C. H., and Speizer, F. E. (1985). Reproducibility and validity of a semiquantitative food frequency questionnaire. *Amer. J. Epidemiol.*, **122**, 51–65.

(CONTINGENCY TABLES
EDF STATISTICS
KAPPA COEFFICIENT
MEASURES OF AGREEMENT
NICKED SQUARE DISTRIBUTION
QUANTILES
TWO-BY-TWO TABLES)

CRAIG B. BORKOWF
MITCHELL H. GAIL

EPIDEMICS AMONG INTRAVENOUS DRUG USERS

The usual model for an epidemic*, as outlined in Bailey [1], assumes that the spread of infection depends on the law of mass action. If, for example, we have a deterministic model in which $x(t)$ susceptibles (individuals liable to infection) mix with $y(t)$ infectives (infected individuals who can transmit the infection) in a population at time $t \geq 0$, then in the absence of a removal mechanism, the number of infectives will increase at a rate proportional to the product $x(t)y(t)$.

While such a model provides a reasonable representation of the spread of infection by contacts between individuals, it is not entirely suitable for infections spread through the exchange of needles among intravenous drug users (IVDUs). Several authors, among them Doll [2], have noted the rapid rise of AIDS cases in the USA and other countries due to the exchange of infected needles among IVDUs. While most studies of epidemics among IVDUs have concentrated on HIV and AIDS, the models outlined below apply to any infection transmitted by blood such as, for example, hepatitis.

Different models have been developed for epidemics among IVDUs, of which two are typical. Firstly, Kaplan [3] has studied a deterministic model based on principles derived from the mass action model. He first finds an equation for the time-dependent probability that an IVDU injects with an infected needle at time $t \geq 0$, and then derives a differential equation involving a mass action factor for the fraction of infected IVDUs at time t in a shooting gallery. He proceeds to analyze various factors such as the sharing rates of injection equipment and their heterogeneity, the mean duration of injection equipment sharing, and the effect of cleaning injection equipment after use. Kaplan provides graphs to illustrate aspects of the model.

Secondly, Gani and Yakowitz [4] have considered a stochastic model where a group of n IVDUs consisting of $X(t)$ susceptibles and $Y(t)$ infectives, with $X(t) + Y(t) = n$, meets regularly at times $t = 0, 1, 2, \ldots$, to inject with drugs. The probability of creating new infectives at such a meeting was derived by Gani [5] in the context of an occupancy problem*. The susceptibles are thought of as cells, while the infective IVDUs who pass their infected needles to them are balls placed in these cells. Assuming the random allocation of infected needles among the susceptibles, and using the occupancy probabilities, one can characterize $X(t)$ as a nonincreasing homogeneous Markov chain whose transition probabilities can be found. It is then easy to calculate the time until the

ENCYCLOPEDIA OF STATISTICAL SCIENCES

UPDATE VOLUME 3

Eds Kotz S, Read C B and Banks D L

A WILEY-INTERSCIENCE PUBLICATION
John Wiley & Sons, Inc. 1999
NEW YORK • CHICHESTER • WEINHEIM • BRISBANE • SINGAPORE • TORONTO